

AN APPROACH TO KNOWLEDGE STRUCTURING
FOR ADVANCED PHASES OF THE
TECHNICAL AND MANAGEMENT INFORMATION SYSTEM (TMIS)

H. T. Goranson, Sirius Incorporated
P.O. Box 9258, Virginia Beach, VA 23452

ABSTRACT: TMIS must employ an enlightened approach to its "object" structure, but basic issues in conceptual structuring remain to be resolved. Sirius outlines the necessary agenda and reports on progress toward solutions.

INTRODUCTION

Sirius is a small group which has traditionally focused on advanced solutions to troublesome issues in knowledge representation. Its prior work has primarily been for larger firms addressing problem spaces in the defense community. Recently, the group has been working in areas directly applicable to the specific problems of object/meta-data/knowledge representation faced by Space Station planners. A relatively mature paradigm is emerging which can be of use in the planning now being done.

THE SPACE STATION INFORMATION SYSTEM (SSIS)

SSIS has embraced the ADA philosophy, which is good. Study groups, particularly the one here at Clear Lake, have developed an agenda and strawman models. Insofar as we at Sirius have seen, the thinking is clear and has led to several early conclusions which we hope will not be diluted as activity increases and more interests become involved. In particular, there is a consensus that a Network Information System (NIS) entity be created and distributed in various incarnations throughout the system to work with (and allow a high level of independence of) "application" systems and communication systems.

Considerations of this NIS are intimately bound with another healthy emerging consensus: before architectures are considered, before coding, module or task planning is begun, even before the development agenda is formalized, a comprehension of the structure of the *representation universe* must be set. This universe must deliberately and circumspectly be engineered as a logical first step in the comprehensive life-cycle planning of the entire system.

This has truly never been faced squarely before, because prior requirements were narrow in comparison and allowed compromise. But SSIS-related issues are too broad, too deep, indeed too expensive *and important* to allow less than our best efforts in this matter.

The problem is: we immediately run up against the same fundamental, unresolved philosophical issues which have plagued the Knowledge Representation community for decades. There is traditional deadlocked contention among

workers in the fields of Artificial Intelligence (AI), Programming Language (PL), and Database (DB) researchers which each represent justifiable perspectives.

The good news is that recent successful studies into the mathematical foundations of allied fields give us a whole new world of tools to work with, tools which allow us to transcend PL/DB/AI localizations and start to put our arms around the science required. This is fully in harmony with the ADA culture, and in fact our group has some ADA-based research underway in this area.

The bad news, at least so far, involves the Technical and Management Information Systems (TMIS), a system planned to support many facets of Space Station development.

TECHNICAL AND MANAGEMENT INFORMATION SYSTEM

While there are major differences between TMIS and SSIS, the TMIS object specification would rightly be placed *before* nearly all SSIS Engineering tasks (Engineering in the ADA sense). TMIS is now in a Phase I ad hoc configuration. Future phases will be competed shortly. What makes the TMIS problem so difficult, and so interesting is the manner in which it must evolve. Each user group will continue to work with his existing and planned environments, constantly adding to the data, information, and knowledge pools.

So the problem, the challenge of TMIS is to develop a conceptual universe which satisfies today's TMIS needs while building the object/representation/meta-data schema, which for simplicity we will call an Integrated Conceptual Environment (ICE). This ICE should be optimized for the emerging ADA centered SSIS, which includes AI interfaces, knowledge representation subsets, and anticipation of future (30 year horizon) technologies.

ISSUES: GENERAL

Basic issues involve ADA language and environment concerns, very large database theory and AI requirements.

Programming Language issues include a complete data and abstract data type specification, associative potential processing topology, as well as many inherent characteristics, such as fault recovery, fault tolerance and fault avoidance. With respect to issues in Programming Language (PL) design, a candidate representational formalism for SSIS characteristics should include constructs that allow a user to capture intuitions about the structure of the domain(s) of application - for example, intuitions about (or clustering operands for) the appropriate conceptualization of the objects, properties, and relations of the domain. The motivation for much of the prior research on incorporating abstract data typing facilities such as ANNA into the ADA programming language has resulted from this incidental requirement. Indeed, in the most general sense, it was this motivation that led to the development of high level programming languages, among them ADA, in the first place.

But neither the syntax nor the "general" semantics of first order languages will recognize the distinction between two types of abstracts, a frequently encountered phenomenon in an anticipated SSIS environment. Things of both types will simply be values of individual variables, and arguments of

predicates (and function-symbols). To some extent, the analogue of this particular semantic inadequacy in programming languages, which is meant to be handled by the introduction of data types and typing facilities, *can* be treated within higher order calculus languages by introducing sortal quantification and by other tedious methods. But this is unacceptable for relating a potentially large abstract vocabulary in a KR context.

Our proposed approach extends a potential semantic capability of ADA into a semantic net operator set, in order to allow creation of a test bed for TMIS development, using current environments.

Concerning DB theory, one of the most striking features of the current explorations in ICEs is the attempt to deploy concepts and techniques from the field of Knowledge Representation, and especially from work on semantic networks, in combination with relational databases. However, one can see something of a paradox in the focus on current network-type formalisms. The backbones of most formalisms are taxonomic hierarchies, sometimes generalized to deal with multiple inheritance. A governing design requirement of TMIS is to capture semantic interdependencies (both inclusions and exclusions) among the various kinds of things (or "concepts") with which they have dealt. But the *explicit* embodiment of such interdependencies seems to be ruled out by the hierarchical relational model.

Given the structural limitations of existing ICE work, it is hard to see how such notions as type hierarchies, inheritance, and exclusion can be imported within the boundaries of the relational model. One proposed handling of hierarchies and inheritance is part of a reconstruction of database theory, from within the theory of first-order logic. TMIS runs the risk of adopting this default position without looking at a higher structural level.

It seems more natural to think about relational databases (especially as conceived of originally) as *models* of kinds of first and second order theories. It is instructional in the present context to propose extending the relational database concept in terms of locally finite dimensional cylindrical algebras. These would be algebras whose elements are sets of finite sequences of objects/attributes, rather than in terms of first or second order theories. We are in fact drawn more to the algebraic, morphological and semantic than to the syntactic, proof-theoretic, account of the relational model in this work. Of course, there are intimate connections between the morphological cylindrical calculus and algebras which our current research exploits, specifically within the unique TMIS universe.

ISSUES: NETS

There has been a long and continuing debate within the AI world on the relative merits and expressive power of representational languages based on Logic on one hand and semantic networks on the other. It is apparent that semantic nets have often been used as a notational rather than as a representational language. Logic, on the other hand, enjoys a well accepted semantics, but its support as a knowledge representation scheme is left unclear as to which aspects of which Logic it proposes to use.

In considering semantic networks, it should be noted that there has been considerable migration of techniques from semantic networks to data models in

the VLD context. Net researchers are just now talking about multiple inheritance, whereas KB researchers have been involved for quite some time. One feature of semantic networks in a purely digital environment that needs to be considered by a data model is the genetics of cluster forces, over a KB life cycle.

A dispassionate observer may question why semantic networks haven't been used more directly for ICE purposes. Semantic networks and semantic data models are in a sense equally powerful, but they have been intended in the past for applications that have different and more linear characteristics. There is a trend in this research to deal with applications that involve an ever increasing number of types, and the gap in intended applications between data models and semantic networks seems to be narrowing. Especially interesting in this case are embryonic studies for photonic machines which appear to be providing a lead in this science of semantic nets.

TMIS will have to consider many possible uses for a database, only some of which would be of relevance to the ICE concerns of SSIS. Their approach is expected to be to determine the *least common denominator* of all these applications and then provide implementation. Although requirements for databases and VLDs have changed during the last few years, it is not obvious that the least common denominators for frameworks that handle photonic forms, visual and speech data, text messages, etc., are semantic networks.

The principals of Sirius, however, represent the philosophy of those who promote the use of nets without the restriction of logic-bound notations, and maintains the pre-eminence of this approach over the other, more constrained historical approaches.

KNOWLEDGE REPRESENTATION (KR) NETS AND LOGIC MATRICES

The theoretical limits of logic-based semantic nets are most clearly seen via properties of topologically isomorphic representations of the data structure. It is easily shown that the generalized map of each is a matrix (for the rule-based "truth table" approach) or a hierarchical structure (for the logic-driven "list processing approach"). The hierarchical geometries are often morphologically classed for queuing purposes and are often related, via pattern recognition layers, as the corners of an all logical space-filling regular three-dimensional tessellation. A commonly encountered example of this method of analysis can be seen in the oct-tree or quad-ternary-like types of machine-level conceptual structures now being investigated by Sirius and other workers in integrated photonics. By examining the morphology of existing Lisp-based methods, it can likewise be demonstrated that a ceiling of complexity results from commonplace simple hierarchical (read "list, stack, or string" processing) building techniques, as expressed in this conventional language syntax.

Consequently, some investigators have applied ingenious techniques to create, access, and manipulate structures of higher morphological complexity. Typical of the approaches is that of Lipski (1978) which advances the available geometry of the data lattices by placing matrix nodes in an ordering which emulates the hierarchical lattice of normal use. This, in fact, parallels a similar synthesis in generating grammars by Paz (1976). In turn, an associative transformational threading is allowed through the lattice, using recursive operators which give the combined structure great conceptual power.

But the theoretical limit on strictly hierarchical geometries remains low in morphological complexity and is limited to those of 3-D, orthogonal, regular tessellation.

Nets of much greater complexity have been used for decades by scientists and mathematicians, especially those involved in work for target image sensor fusion. Recently, the regular, infinite structures that are generatable by periodic recursion have been categorized by Goranson (1981), Burt/Wachman (1974), and Lalvani (1986). The approach to establishing a baseline vocabulary for the ICE comparative optimization studies selects a structure for the knowledge representation lattice based on the desired properties of the task and nature of the diversity of the source data and internal translative types.

This "Top Down" approach gives many advantages relevant to related *connection* biased theories. Major advantages for this proposed approach include related reduction of a meta-operator set (using primitive and recursive transforms) to a few mathematical symmetry operations to which we can apply several optimization techniques. Greatly increased density and organization of the information in the abstracted knowledge bases is achieved. This latter advantage allows a suitable level of abstraction to be combined with clustering of related elements (which can be seen as extended "frames"), a desirable effect which addresses suitability for local TMIS users. Memory requirements are greatly reduced in size. Elements which express cluster characteristics are typically quite large in number when narrow in scope and are well ordered for logical calculus interface. The resultant partitions are attractive for concurrent and parallel processing algorithms on a level low enough to include technologies projected for SSIS use.

IMPLICATIONS FOR THE PRESENT PROBLEM

The requirement is for an approach which:

- a. Will provide a basis for synthetic studies of applicable technologies and methods across a wide spectrum, in short: a descriptive ICE nomenclature.
- b. By virtue of providing a high level scientific description of the laws at work, will provide a unified means for integrating this diversity.
- c. Allow an approach to optimizing ICE specification, design and hosting, provide for future generation computing requirements in terms of processing, large information sources, and reconfigurability.
- d. Direct specification for hosting and testing of these issues.

GENERAL TOOLS

Sirius works with a set of morphological expressive laws which describe all possible global model schema. This morphology is currently embedded in a set of programs developed over time and currently used for internal research and is strictly within the ADA philosophy.

A specific tool cited, ALICE, described here as central to the approach is a method of machine hosting ICE morphological operators in ADA. This allows any high level semantic net to be quickly defined and tessellated among attributed space. While the geometries involved originate in Euclidean lattices, restrictions on three dimensional analogues disappear so the number of dimensions addressed is limited only by practical considerations of the size of the symmetry tensor used. Because TMIS requires operations that do not exclude direct microcoding in embedded systems, the number of dimensions is practically limited to 16 based on anticipated hardware, assuming that the lattice chosen has less than 3 symmetry characteristics. (Some work has been done by Sirius using linear infinite polyhedral geometries, hosted on fine-grained arrays, but the hardware required is specialized, and beyond the scope of this TMIS work.)

Major VLD and DB investigative interface tools are DAVID (from Goddard) and TIPS (from DoD). For determining the lattice geometries, the work of Lalvani and Burt forms the basis for long term studies by Sirius, and are employed here, following the symmetry operators proposed by Schoen (1970) and Wells (1977). The basic research of Brisson is the source for the multi-dimensional holomorphic linearizations required, as well as a source of understanding higher dimensional clustering effects when mapped to lower level matrices, which in turn follows a suggestion of Williams (1969).

Key to the approach is the intersymmetry minimal surface operator, an approach devised by Burt for the generic translation of intersymmetry groups. A complementary approach utilizing added dimensions of symmetry variables is used to annotate the entity attributes/relationship attribute from Lalvani.

The minimal surface operator methods themselves follow closely the trends established by Jenkins (1966), after Blatter (1971) and Voss. The actual trace algorithms were suggested and bounded by Rahimi (1972) and Barber (1970). The reflexive nature of that algorithm which reflects the Lalvani transpolyhedra loosely follows the technique applied by Yoshizawa (1982) to a similar problem in dynamics. Provisions for fractal doubt, not discussed here, come straight "out of the book" from McClure, as well as a few other specialists, not currently in the free world.

Lacking full-time access to supercomputer facilities, the present capability for VLD simulation and storage of multidimensional matrices that Sirius possesses is inadequate. A satisfactory, and expedient cost effective solution to this inadequacy has been found in the N-Dimensional Data Base System. This software exploits a special case of the symmetry storage case, namely the Generalized Balanced Ternary. It has been determined that the restrictions imposed on VLD research are far outweighed by the availability of this tool for the early phases of the effort. Rehosting overlays to TMIS host systems should be straight forward.

OVERVIEW OF ALICE

Sirius has been working this problem for some time. As described, a high-level conceptual language is required, for this type of effort reflecting the ponderous capabilities implicit in the problem. The result of Sirius' work is ALICE (ADA/Lattice Integrated Conceptual Environment). ALICE is coded in ADA, of course, and uses as operators, translation sets which are created by the abstraction tools described.

An outline of the features of ALICE:

- a. The relational function set is extended into a primitive operator language. This language has a base logic of some order and has a mapping to the network lattice transforms.
- b. These transforms have topological equivalences, a unique and interesting feature which allows the rules for operation among the fabric to be ambiguous when mapping "up", but precise when mapping "down", or out of abstracted spaces.
- c. The topological equivalence also allows resolution into a few simple operations (which include photonic operators) which can be matrixed and solved by arithmetic array processors. This holds promise for high speed, concurrent AI processing using VHSIC; Reduced Environment Math processor arrays.

Having established a meta-language with a corresponding meta-net, and formal methodology the ICE researcher is faced with several exciting possibilities. Knowledge of any kind can be "fused" by simple procedures if the source calculus is descended from the universal structure employed.

For example, image data, graphic modeling information, engineering data (which may not necessarily be geometry dependent) and performance information can be all imposed on the same ICE. As relationships are established, the diverse information becomes fused, and is actually aggregated by virtual proximity; *Virtual Proximity* is defined as relation by a critical set of algebraic operators defined by the sliding threshold factors of the periodic lattice matrix.

BIBLIOGRAPHY

Lipski, W., "On Semantic Issues Connected with Incomplete Information Bases", TODS 4 (3), 1978.

Paz Azaria, "Multidimensional Parallel Rewriting Systems:", in "Automata, Languages, Development", ed Aristid Lindenmayer and Grzegorz Rosenberg, North Holland, 1976.

Goranson, H.T., "Periodic Recursion Dynamics in Large Integrated Conceptual Environments", Sirius, 1981

Wachman, A., M. Burt, and M. Kleinman, "Infinite Polyhedra", Technion, Haifa, Israel, 1974.

Lalvani, Haresh, "Structures on Hyperstructures", Privately Published, 1982.

Lalvani, Haresh, "Transpolyhedra", Privately Published, 1977.

Schoen, Alan H., "Infinite Periodic Minimal Surfaces Without Self Intersections", NASA TN D-5541, 1970.

Schoen, A.H., "Infinite Quasi-Regular Warped Polyhedra and Skewness of Regular Polygons", Not. Amer. Math. Soc., Vol. 15, 1968, pp 801.

Wells, A.F., "Three-Dimensional Nets and Polyhedra", (Chapters 19 & 20), Wiley, Interscience, 1977.

Brisson, David W., "A Definition of Distance and Method of Making Space-Time Measurements", Speculations in Science and Technology, Vol, 3, #5, 1980 pp 579-586.

Williams, Robert Edward, "Dimension as Level" in "Hierarchical Structures", ed. Whyte, Wilson and Wilson, American Elservier, 1969.

Jenkins, Howard and James Serrin, "Variational Problems of Minimal Surface Type III, the Dirichlet Problem with Infinite Data" Archive for Rational Mechanics and Analysis, Vol 21, #4, 1966.

Blatter, J., and E.W. Cheney, "Minimal Projections of Hyperplanes in Sequence Spaces" AFOSR-2130-71, Report NTIS AD 735 759 1971.

Rahimi, A. and R.W. Brockett, "Homotopic Classification of Minimal Periodic Realizations of Stationary Weighting Problems", SIAM Jour Applied Math 22, #3, pp 481-486, 1972.

Barber, M.N. and B.W. Ninham, "Random and Restricted Walks: Theory and Applications", Gordon and Breach, 1970.

Yoshizawa, T., "Stability Theory and the Existence of Periodic Solutions and Almost Periodic Solutions", Applied Mathematical Sciences #14, Springer-Verlag.

McClure, Donald E., "Image Models in Pattern Theory", Computer Graphics and Image Processing, Vol. 12, pp 309-325, 1980.